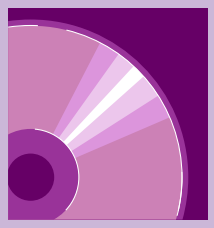
Using the 1990–1993 Full Panel Files



This section focuses on information specific to the full panel files.

Because the 1996 full panel file is not yet available, the information in this section applies only to the 1990–1993 full panel files.

Structure of the 1990–1993 Full Panel Files

The full panel files contain one record for each person who was ever in the SIPP sample for that panel. This is true even if the person was in the sample for just 1 month. Full panel files contain records for children and for people who entered the sample after the first wave.

Within each record, variables correspond to the information collected in the core interviews. However, some core items, including some constructed variables, are not included on the full panel files. No items from the topical module files are on the full panel files. tip

SIPP tip

Analysts familiar with the core wave files should be careful when using the full panel files. Important information about families, unearned income, and other key topics is coded and/or organized differently in the two file types.

Using the Data Dictionary

The format of the data dictionary for the 1990–1993 full panel files is similar to that used for the pre-1996 core wave and topical module files except that two extra fields are added to lines with a "D" in the first column. These two fields denote:

- The number of occurrences of the variable (for example, some questions were asked each wave of the panel, and some questions were asked each month of the panel)
- The number of digits for each occurrence of the variable tip

tip

The data dictionary for the 1992 full panel file has a line labeled with an "R" in column 1. This line provides value ranges for the variable. Also, fields in lines beginning with a "D" vary somewhat from "D" fields in other full panel files.

Aligning Data by Calendar Month

Analysts often find it useful to realign SIPP data by calendar month rather than reference month. For example, to analyze data for a specific calendar year or fiscal year, SIPP users must realign the data.

There are various approaches for realignment. In each case, the analyst must use the technical documentation to determine the reference period for each rotation group of the panel. Analysts also need to apply the mapping from reference month to calendar month for each person included in the analysis.

Chapter 12 of the *SIPP Users' Guide* contains an algorithm that realigns data by calendar month. In the algorithm, the first step realigns the months; the second step initializes each monthly variable to distinguish the months in which the variable is not relevant. Finally, the algorithm realigns the input data to be based on the calendar month.

Link to the algorithm.

Identification/Description Variables

Monthly Interview Status

In the full panel files, the monthly interview status variable (PP-MIS), which helps determine whether data for a person in a given month should be used, occurs once for each reference month of the panel. Analysts should use data only for months in which the interview status variable has a value of 1. tip

SIPP *tip*

Analysts should be careful not to confuse the monthly interview status variable with the interview status variable (PP-INTVW).

Identifying Persons

To uniquely identify a person in the 1990–1993 full panel files, analysts should use the following three variables:

Variable Name	Description
PP-ID	Sample unit ID
PP-ENTRY	Entry address ID
PP-PNUM	Person number

PP-ID is a random recode of three variables in the Census Bureau's internal files. The variables are omitted from the public use files to protect the confidentiality of respondents.

Identifying Households

To uniquely identify households and group quarters in the 1990–1993 full panel files, analysts should use the following variables:

Variable Name	Description
PP-ID	Sample unit ID
HH-ADDID$_i$	Current address ID in the *i*th month
PP-MIS$_i$	Person's interview status in the *i*th month

Because household composition changes from one month to the next, it is generally not possible to construct "longitudinal households." For a given person, analysts should evaluate the characteristics of the household each month. Characteristics should cover only those people who reside together in each specific month.

Identifying Families

Unlike the core wave files for the 1990–1993 Panels, the corresponding full panel files do not contain family identification variables (e.g., FID, FID2, and SID). Analysts needing family identification variables must either merge them from the core wave files or create them. Because family composition can change over time, these are monthly variables. tip

Link to an algorithm that provides one approach to creating functional equivalents of the variables on the core wave files.

Describing Family and Household Composition

Analysts can use the household ID variables and the variables created by the "family" algorithm to group people into the same family and subfamily groups that appear in the core wave files. However, the actual values assigned by this algorithm to these variables generally will not equal the values found in the variables from the core wave files.

The 1990–1993 full panel files also include nine additional variables that can be used to identify relationships to reference persons and a variety of household configurations, including households containing three generations.

Link to a table containing the nine household description variables.

Identifying Program Units

The 1990–1993 full panel file information on participation in health insurance and government transfer programs differs in some ways from the corresponding core wave file information.

SIPP*tip*

Beginning with the 1991 Panel, a new missing wave imputation procedure was applied to the full panel files: data were imputed for people with missing data for a wave but with valid data for the two adjacent waves. For these people, merging the core wave family ID variables is not an option.

1. In the full panel files, the authorized recipient variables do not use the entry address and person number values. Instead, they use the sequence number of the person within the sample unit (PP-RCSEQ) to identify authorized recipients. For example, the authorized food stamp recipient is the person for whom FS-PIDXi in month i equals PP-RCSEQ.
2. The variables used to identify members of a common program unit in a given month (i) can be identified with the following three variables:
    - Sample unit ID—PP-ID
    - Person's interview status in month *i* —PP-MIS$_i$
    - Authorized recipient variable in month *i*
3. Unlike the core wave files, the full panel files have no coverage variable indicating whether the child, adult, or both were covered by SSI. If needed, that information can be acquired from merges with the core wave files.

Identifying Movers and Household Composition Changes

The procedures for identifying movers and household changes are essentially the same in the 1990–1993 full panel files as in the corresponding core wave and topical module files. In the rare cases of persons in merged households who were assigned new ID values, the full panel files contain two full panel records for those persons.

Chapter 12 of the *SIPP Users' Guide* describes the procedures for tracking movers in the 1990–1993 full panel files.

Identifying States and Metropolitan Areas

States. SIPP is not designed to allow analysts to produce state-level estimates. A state variable (GEO-STE) is included on the 1990–1993 full panel files to allow examination of how state-level estimates affect national-level estimates. GEO-STE identifies 41 individual states and the District of Columbia; the remaining 9 states are suppressed into three groups.

A user could apply the state-specific eligibility criteria for a means-tested program to arrive at a national estimate of the number of people eligible for the program.

Metropolitan Areas. The full panel files do not contain any variables identifying metropolitan areas. Analysts needing that information must merge it from the core wave files.

Income Variables

Family-Level Income Variables

The family-level income variables in the full panel files, like those in the core wave files, include the income of all related subfamily members. However, unlike the core wave files, the full panel files do not contain any subfamily income variables. If family income variables are needed that do not pool related subfamilies with primary families, those income variables must be created. tip

SIPP *tip*

Unpooled income variables can be created by looping over persons with PP-MIS$_j$ of 1 and with common PP-ID, HH-ADDID$_j$, FID2, and SID$_j$ for each month.

Unearned Income Variables

Analysts need to be aware that the Census Bureau organizes the unearned income variables differently in the core wave and full panel files.

In the full panel files, 10 variables on each person's record identify up to 10 different sources of unearned income. For each source identified, there is a corresponding amount variable.

When using the unearned income fields in the full panel files, analysts often find it helpful to realign the unearned income into new income-specific variables.

Link to an algorithm that demonstrates how to create income-specific variables.

Income Topcoding

Income topcoding procedures in the 1990–1993 full panel files are the same as those used in the core wave files of the 1990–1993 Panels.

Using Allocation (Imputation) Flags

The edit and imputation procedures used for the 1990–1993 full panel files differ from those used for the corresponding core wave files. The procedures for the full panel files make use of a full set of longitudinal data for a person, in contrast to a maximum of 4 months of observations that can be applied to the core wave files. The procedures applied to the core wave files make greater use of cross-observation imputation methods than do those applied to the full panel files. tip

Two sources identify whether information has been imputed in the 1990–1993 full panel files:

1. Beginning with the 1991 Panel, all data for a wave are imputed if a person was not successfully interviewed in one wave but had complete information (from either a successful interview or a proxy interview) in the two adjacent waves. In those cases, the value of WAVFLG will be greater than zero and INTVW will be 3 or 4.
2. Imputation flags appear for a limited set of variables, including earned income, asset income, and unearned (transfer) income variables.

Weight Variables

The 1990–1993 full panel files include:

- The calendar year weights—FNLWGTs
- The full panel weight—PNLWGT

The number of calendar year weights corresponds to the duration of the panel.

SIPP *tip*

The edit and imputation procedures applied to the core wave files from the 1996 Panel make greater use of prior wave information than procedures used in earlier panels.